Warwick Economics Research Paper Series

Competence vs. Loyalty: Political survival and electoral fraud in Russia's regions 2000–2012

Christoph Koenig

November, 2015
Series Number: 1080
ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Competence vs. Loyalty: Political survival and electoral fraud in Russia's regions 2000–2012*

Christoph Koenig

The University of Warwick†

November 12, 2015

Latest version: christoph-koenig.org/russianelections.pdf

Abstract

Election fraud is a pervasive feature of autocracies but often only serves lower-tier officials to cast signals of loyalty or competence to the central government in order to pursue their own interests. How much such personal interests matter for electoral fraud under autocracy has however not been studied so far. In this paper, I exploit a radical policy change in Russia which allowed the president to replace governors of the country's 89 regions at his own will. As a result, federal elections after December 2004 were organised by two types of governors: one was *handpicked* by the president, the other one *elected* before the law change and re-appointed. Even though both types faced removal in case of bad results, the need to signal loyalty was much lower for the first type. In order to estimate the effect of handpicked governors on electoral fraud, I use a diff-in-diff framework over 7 federal elections between 2000 and 2012. For this time period, I use results from about 95,000 voting stations to construct a new indicator of suspicious votes for each region and election. I show that this indicator correlates strongly with incidents of reported fraud. My baseline estimates show that in territories with a handpicked governor the share of suspicious votes decreased on average by more than 10 percentage points and dropped even further if the region's economy had done well over the past legislature. These findings suggest that governors have less need to use rigging as a signal once loyalty is assured unless faced with circumstances raising doubts about their competence.

*I would like to thank Sascha O. Becker, Peter Buisseret, Ernesto Dal Bó, Mirko Draca, Ruben Enikolopov, Mark Harrison, Giovanni Mastrobuoni, Sharun Mukand, Maria Petrova, Kirill Pogorelskiy, and Fabian Waldinger for important discussions and useful insights and Maxim Krukov of GOLOS and Graeme Robertson for sharing their data. The paper further benefited from seminars and participants at Warwick. Financial support by the ESRC DTC Warwick is gratefully acknowledged. All remaining errors are mine.

†Department of Economics. Email: C.Koenig@warwick.ac.uk

> I think you need to pay attention to those areas where our people have denied United Russia serious trust. Not because it is a tragedy, but because it is a signal for the authorities.
>
> Dimitri Medvedev, 2011

1 Introduction

One common explanation for the poor economic performance of autocracies is the failure to hold leaders accountable for bad policy. While democracies use free and fair elections in order to punish or reward politicians, the voting process under dictatorship is often manipulated and seen as a meaningless political ritual. Recent literature has called this view into question arguing that authoritarian elections are actually used by the ruling circle to hold officials accountable. Similar to the role of education in the labour market, elections may be used by subordinates to cast signals about their loyalty or competence to their superiors (Gandhi and Lust-Okar, 2009). For instance, high results may be rewarded by directing additional resources under the control of the respective official or by advancing his party career chances through promotion (Martinez-Bravo, 2014). With legal barriers largely absent, this creates strong incentives for public officials to artificially change results and engage in electoral fraud. The extent to which such concerns are driving ballot rigging in autocracies has, however, remained a largely understudied topic since institutional setups under dictatorship tend to be rigid, and reliable data on the intensity of fraud is scarce.

This paper uses detailed data on Federal elections in Russia and a novel tool for detecting suspicious results to study the effect of a radical policy change in accountability to the central government. After the 2004 Beslan hostage crisis, which exposed severe inefficiencies in the local administration, president Putin signed a law which abolished governor elections in the country's 89 regions. From December 2004 onwards, regional leaders thus had to be appointed by the president in order to stay in office after their term but could also be dismissed without any legal barriers (Hill, 2012). While this gave the central government an important *stick* to punish notoriously under-performing and corrupt leaders, it also severely altered their motivation to please their superiors by delivering the *right* results in federal elections which they can organise in their territory at large discretion. The reality of

this threat can be seen from the low election results of United Russia in December 2011 which resulted in the dismissal or *voluntary resignation* of 6 governors over the following 4 months (Chaykovskaya, 2011; Moraski and Reisinger, 2013).

The legal change, however, also started a transition period from 2005 onwards during which *handpicked* and *elected* governors coexisted and organised 4 national ballots. This allows me to study the differential effect of abolishing governor elections on two types of leaders: one whose loyalty is assured but uncertainty about competence may still exist, and another one where neither of these is known (see also Egorov and Sonin, 2011). In my conceptual framework I argue that, for given levels competence, appointments may induce *lower* levels of election fraud since officials do not need to use rigging as a means to signal their loyalty to the centre. In a recent paper, Martinez-Bravo (2014) studied the effect of local officials on election fraud in Indonesia where a democracy inherits a set of potentially not trustworthy officials from an autocracy. The author demonstrates empirically that appointed village heads have a higher need to convince their supervisors of their suitability than elected ones. The mechanism I propose is dealing with the opposite problem faced by an authoritarian government: until all positions are filled with loyal party cadres, officials chosen under the previous regime are particularly suspicious of disloyalty and may therefore try to over-compensate. Moreover, my model claims that, when facing a bad economic performance, also leaders selected by the central government may engage in fraud to send a signal about their competence and keep their position.

In order to empirically test my predictions and study the effect of this policy change on election fraud, I use unique highly disaggregated data at the voting station level for all 7 national elections – parliamentary and presidential – held in the Russian Federation from 2000 to 2012. For each region I calculate the percentage of votes cast in districts with highly suspicious results. A district's result is deemed suspicious if the turnout and vote share regression coefficient (TVSC) takes values greater or equal to one. The TVSC has been widely used in the study of electoral fraud in the Russian context and is appropriate for detecting ballot stuffing and other turnout-inflating types of manipulation (Myagkov et al., 2009). Legitimate doubts about its reliability and shortcomings are met by a number of tests. First, I show that my measure is significantly correlated with reported incidents of fraud during the 2011 and 2012 elections and decreases with the introduction of electronic vote scanners across Russian regions. Second, I benchmark the TVSC against other indicators using first- and second digit distributions of incumbent vote and valid ballot totals which do not seem to have similar power in detecting rigging in Russian elections.

I quantify the treatment effect of governor selection using a differences-in-differences estimation which controls for region-specific and election-specific determinants of rigging and the time-varying effect of pre-2000 democratic institutions. Competence of a governor is measured by an index of regions' changes in unemployment and GDP per capita growth over the last legislature of either president of parliament. Both variables are then aggregated to a single index of economic performance over the last 4 years. The regression specification includes this index as well as its interaction with the treatment in order to study the differential response to economic fluctuations across both handpicked and elected governors. The baseline results show that regions with a handpicked governor obtain on average more than 10% less votes from districts with highly suspicious results. Second, the negative effect of a handpicked governor on fraud is even stronger during times of good economic development. Only under extremely bad conditions does the negative marginal effect of a handpicked governor disappear completely. I use various checks to address concerns about the endogeneity of governor replacement such as placebo tests for different time periods and outcome variables. The coefficients are robust to the inclusion of region-specific time trends and election fixed effects for each Federal district as well as different definitions of economic performance. Furthermore, placebo experiments show no effect if treatment is moved one or two elections forward and no response on unrelated election outcomes such as votes for other parties.

Several empirical studies have investigated the fate of Russia's governors after 2005 but mainly focussed on the selection mechanism (Reuter and Robertson, 2012; Moraski and Reisinger, 2013) or outcomes other than election fraud (Moraski and Reisinger, 2009; Rochlitz, 2013)[1]. Kalinin and Mebane (2011) are studying how federal transfers affected election fraud in the 1990s and 2000s using a number based approach and aggregated data. I contribute to this literature by linking the incentive structures of governors after December 2004 to changes in election fraud over time. The TVSC method for detecting fraud used in this study has furthermore been used in various papers on the Russian context (Filippov and Ordeshook, 1997; Myagkov et al., 2009; Lukinova et al., 2011; Enikolopov et al., 2013), yet so far without a systematic application over all regions and several elections. My paper adds to this work by providing further evidence on the TVSC's ability to capture fraud and by making it usable for cross-regional comparison over time. Studies of fraud have been carried out for other countries using different indicators of ballot rigging. Examples of these are Ziblatt (2009) for Imperial Germany, Ichino and Schündeln (2012) for Ghana, and Cantu and Saiegh (2011) for Argentina. With the exception of Ziblatt

[1] See Rochlitz (2013) for further references on this topic.

(2009), my work is one of few papers studying the evolution of electoral fraud over time. In terms of the studied mechanism, the closest piece of work is the previously mentioned study by Martinez-Bravo (2014) who looks at the topic from the side of a democracy and proxies election fraud through high votes for the dominating party instead of a direct indicator of rigging.

The paper starts with description of important institutional details and the conceptual framework used to analyse the changing incentive structures of election fraud in Russia. After briefly describing the data used, I present the main fraud indicator with a special emphasis on its reliability and comparison with other potential alternatives. Next, I outline the differences-in-differences approach used in the empirical analysis and discuss the validity of its assumptions in the studied context. The baseline results are presented in the subsequent section which is followed by robustness checks and placebo tests. The last section concludes.

2 Institutional and theoretical background

2.1 Relevant aspects of Russia's political system

As in any presidential system, the president is paramount in Russia's constitution. He appoints the government as well as the members of the constitutional court and the supreme court. In addition, he has the right to veto laws passed by the legislative and can also initiate laws himself (Chaisty, 2012). Moreover, he can dissolve the State Duma under extreme circumstances, rule by decrees without consent of the parliament and call for an emergency state which gives him the power to even ignore civil freedoms. The president is chosen in national elections and the length of term has recently been extended from four to six years from 2012 onwards (Sakwa, 2008). The legislature of the Russian Federation consists of two chambers - the State Duma and the Federal Council. The main task of the Federal Council is to represent the Russian regions. It has 178 members, with two representatives for each administrative unit. Representatives are appointed by the regional executive and the regional parliament and can be withdrawn by these institutions. The State Duma's 450 members, in turn, are elected through a national ballot. Until 2007, half of the deputies were elected by majority in single-member districts and the other half proportionally through party lists. This system was, however, abandoned in favour of a purely proportional representation. Like the president, both chambers have the power to initiate new laws (Sakwa, 2008).

As in most federal states, each subdivision has its own legislative and executive. Each Russian region is headed by a governor.[2] The members of the regional assemblies are chosen in local elections. Similar to the national level, these assemblies are dominated by the executive making governors the most important political institution in the regions. From 1996 until December 2004 they were chosen in local elections. However, in 2004 the constitution was changed in favour of a direct appointment of governors by the president (Slider, 2012). This drastic constitutional change was decided in the aftermath of the Beslan Massacre. On the 1st of September 2004, a multinational terror squad took over 1000 hostages in a school in Beslan, a town in the Republic North Ossetia-Alania close to Chechnya. When security forces attempted to free the hostages, more than 300 people were killed.

This national tragedy demonstrated the increased power of Chechen insurgents and their allies but also showed the lack of coordination between federal and regional authorities.[3] Very soon after the attacks, president Vladimir Putin initiated a law which re-introduced the appointment of governors. The draft passed both chambers of the Federal Assembly and came into effect in December 2004. What may seem puzzling is that the new law was accepted by both the population and the governors without any major opposition. Goode (2007), who analysed the parliamentary debates in late 2004, concludes that a combination of rally-around-the-flag effects and an appeal to Soviet legacies made it impossible to reject the new law. Additionally, being independent of the local electorate and depending only on the central executive was in the interest of many governors.

2.2 Conceptual framework

Election fraud in this paper is perceived as the result of an interaction between the president P and a governor. The latter can be of two types, either elected (G_E) or handpicked (G_H). Governors organise national elections in which P runs for office and can exert fraud in order to influence the results in P's favour. P's stay in power is assumed not to depend on the election outcome but he cares about the governor's competence C and loyalty L.[4] Competence is appreciated by P for reasons uncorrelated with election outcomes such as international reputation or development assistance. In line with Egorov and Sonin (2011) I assume, however,

[2] Many regions use different titles such as *President*, *Head of the Republic*, or *Head of the administration*. For the sake of simplicity, I refer to all these in this paper simply as *governors*.

[3] This was apparent even though a lot of information about the Beslan hostage crisis was actually withheld from the Russian public (Haraszti, 2004).

[4] One could imagine, for instance, that centralised state propaganda ensures high levels of political support for the incumbent.

that L matters far more than C for P. While L is known by P, he uses election results to infer the competence $\hat{C}$. The interaction between P and G can be separated into two phases. In the first one, G_E is elected by popular vote and therefore L is unknown. For simplicity, I assume that L_E is 0. Since fraud is a costly action and P cannot hold G_E accountable, there is no fraud in this phase.

In the second phase, P is now equipped with the power to remove an elected governor G_E. P's decision is based on whether it is more beneficial to select a new governor G_H from his cadres who is loyal with probability 1 but of uncertain competence. This decision is based on evaluating the last election results. For simplification, I make the strong assumption that P is not well informed about voters' decision making process. In fact, he interprets his vote share as a linearly increasing function of $\hat{C}$. Governors can, however, exert fraud in order to artificially increase the result and influence P's conclusion about $\hat{C}$. P's decision about keeping or dismissing governors depends on how their loyalty and competence compares with that of a handpicked new governor. If $\hat{C}$ is below of the minimum required competence $\hat{C}^{min}$, he will be removed. As a result, minimum competence levels will vary by type such that $\hat{C}_E^{min} > \hat{C}_S^{min}$. Hence, the only way for G_E to compensate his lack of loyalty and stay in office is through artificially increasing $\hat{C}$ by means of electoral fraud which motivates the first hypothesis.

Hypothesis 1: *For given levels of competence, handpicked governors will engage in less rigging than elected ones.*

The second difference between the two types of governors is the way they are evaluated by the electorate. Unlike his elected counterpart, G_H knows that his own economic performance has an influence on votes for P. The reason for this is, that voters may want to punish or reward P for his choice of governor.[5] Assuming that economic performance is stochastic and independent of C to some degree, it follows that if a handpicked governor does a bad job, he is more likely to be fired. This is because P will infer a lower quality, unless G_H compensates this through rigging. Vice versa, a good performance allows him to reduce the amount of fraud even further. The second hypothesis therefore goes as follows:

[5] That such behaviour is actually at work is well exemplified in the study by Szakonyi (2012) who investigates the political reactions to the wildfires in Russia in 2010.

Hypothesis 2: *Handpicked governors will engage in less rigging if they performed well during the last election period.*

Having presented the main hypotheses to be tested in this paper, I now proceed to discuss my measure of election fraud used in the empirical analysis.

3 Measuring election fraud

3.1 The turnout/vote share correlation

The turnout/vote share indicator was first applied by Sobyanin for the 1993 constitutional referendum and is probably the most widely used tool for detecting election rigging in Russia. It is most suited for turnout-inflating cases of fraud and relies on the assumption that within a given entity and absent manipulation there should be no correlation between how many people vote and their choice across lower-tier areas. Figure 1 illustrates this with a brief example similar to Myagkov et al. (2009): there are 24 voting stations in an area with a given homogenous support of 75% for candidate i. Half of the stations are in high-turnout areas where 60% of the electorate casts their ballot, whereas the remaining ones only have a turnout of 40%. Absent fraud, a 1% higher turnout T is associated with an increase of 0.75% in votes for i out of the total electorate, V/E. A simple OLS regression thus yields a turnout/vote-share coefficient (henceforth $TVSC$) equal to the average support of the candidate. This relation, however, would not hold in the case of ballot stuffing or other turnout inflating methods of manipulating the outcome as can be seen from the right panel in figure 1. In this scenario eight of the formerly low-turnout stations see their turnout artificially increased with all additional votes going to candidate i. The TVSC thus changes from 0.75 to 1.07 which cannot be equal to i's natural support in that area anymore.[6]

Following Myagkov et al. (2009), one can distinguish between the cases when 1) the TVSC exceeds the candidates vote share in the respective area but is smaller than one and 2) the TVSC is bigger or equal to one. In the first scenario, the conclusion is ambiguous and will only be a safe detector if one can rule out that the favoured candidate – absent fraud – would have fared particularly well in lower-tier areas of high turnout – a premise which is quite difficult to check. TVSC ≥ 1 appears

[6] One could imagine a scenario in which fraud is conducted in such a way that turnout and vote share are identical in each voting station. In this case there would be no variation and a TVSC could not be calculated. While this is theoretically possible, it is very difficult to implement in reality. I did not encounter such a case during the construction of my fraud data.

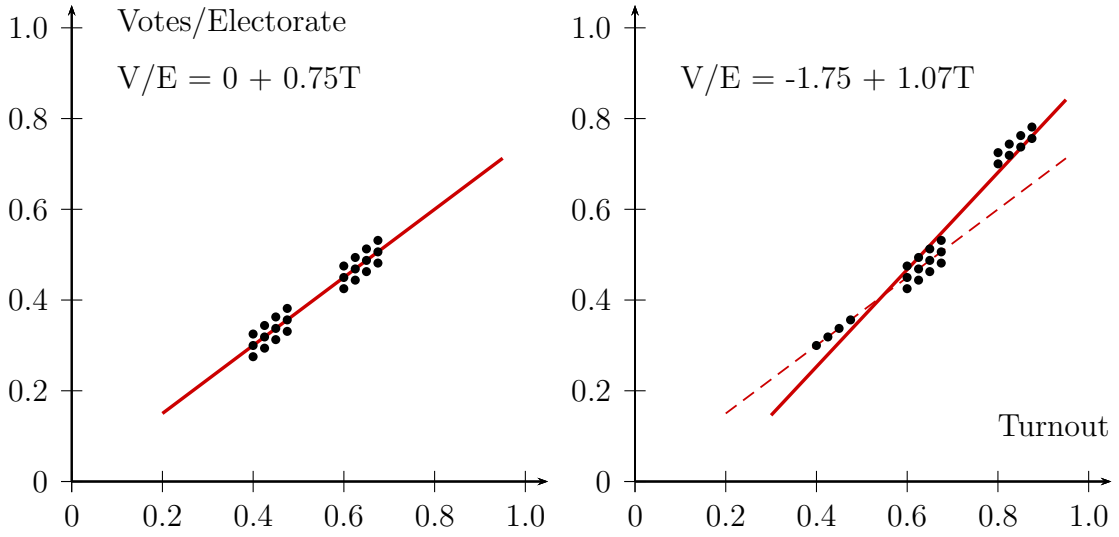


Figure 1: Example of TVSC absent fraud (left) and present (right)

to be a stronger indicator of manipulative turnout inflation, but it is also prone to fallacies as depicted in figure 2. One may think of a region with uniform support of 75% for candidate i across its three districts with 4, 12, and 8 voting stations respectively. The voting stations, however, are not homogeneous since districts differ substantially in their average turnout level. A regression of V/E on T across the whole region will thus suffer from aggregation bias and yield a TVSC larger than one even though fraud did not take place. While this error cannot be ruled out entirely, it can be mitigated by using highly disaggregated data and calculating the TVSC for reasonably homogeneous areas. Enikolopov et al. (2013), for instance, have shown that the random allocation of election observers across voting stations in the city of Moscow during the 2011 Duma election significantly decreased the TVSC calculated for the United Russia party.

Further estimates of election fraud in Russia's regions using the TVSC have been scarce so far and mostly relied on district aggregates (e.g. Myagkov et al., 2009). In these cases the assumption of homogeneity is difficult to defend and the amount of districts/observations may be very low.[7] The availability of election results at the voting stations level since 2000 allow me to calculate TVSCs in each district of a given region (e.g. Lukinova et al., 2011) and to construct new and more robust estimates of election rigging across Russia's regions. As a new measure of regional fraud intensity, I propose the *share of votes from districts with a TVSC* ≥ 1. This indicator has the main advantage of using data from comparatively small areas such

[7] The city of Moscow, for instance, has 130 TIKs while the Nenets Autonomous Okrug has only three.

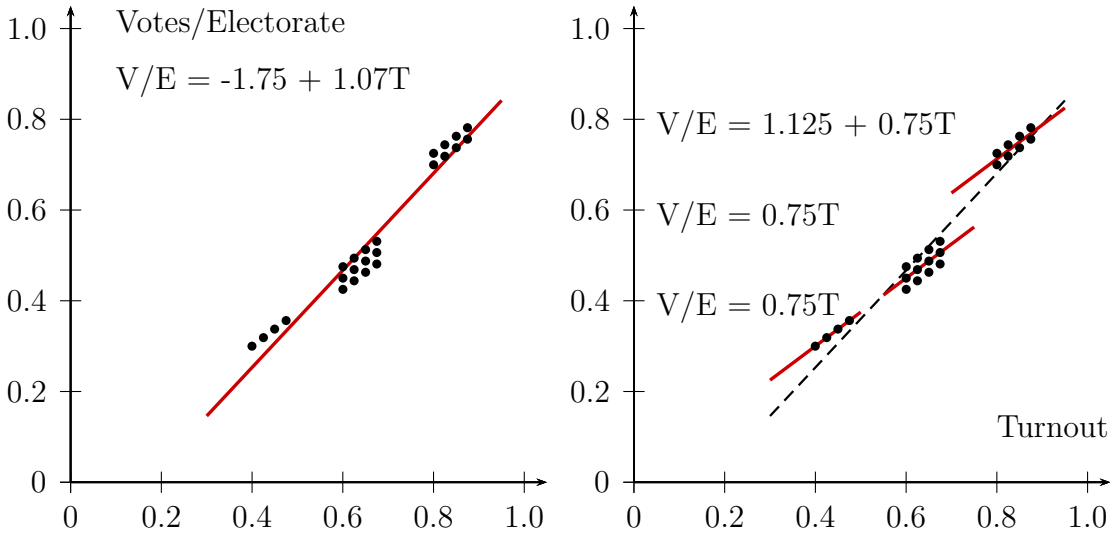


FIGURE 2: EXAMPLE OF BIASED TVSC IN THE CASE OF NON-HOMOGENEOUS AREAS

as districts but simultaneously provides a regional aggregate from this information. It is supposed to capture the intensity of rigging rather than its mere existence which has been observed in virtually every region across the country and therefore does not provide much information. Another interpretation of the indicator is the percentage of votes likely to be affected by manipulation or, simply, the *suspicious vote share*.

The choice of this indicator appears arbitrary at first and one may think of several alternatives: the first option could be to deem a district's votes suspicious if the TVSC exceeds the candidate's vote share and calculate the suspicious vote share based on this rule. Alternatively, one could estimate a single region-specific TVSC from all voting stations and calculate either a dummy for whether it exceeds one or obtain the continuous difference from the candidate's actual vote share. The following section compares the proposed indicator of fraud to these alternatives in terms of reliability and provides further checks of its validity.

3.2 Reliability and validity checks

Before proceeding with a specific indicator of election fraud, one needs to assure that it is reliable and valid. Verifying reliability beyond anecdotal evidence is particularly difficult in the context of election rigging since officials usually try to hide their actions. I tackle this issue with fraud report data from the NGO *GOLOS* (Russian for *vote* or *voice*). This provides information on the region where election irregularities were witnessed. Subsequently, one can form a regional measure and

Table 1: Reported irregularities 2011-2012 and fraud indicators based on TVSC

Reports per 100k electorate	Effect of TVSC based on			
	vote share, with TVSC		regional aggregate, with TVSC	
	≥ 1	$\geq$ Incumbent vote	≥ 1	$\geq$ Incumbent vote
	(1)	(2)	(3)	(4)
Improper counting	0.547***	0.224	0.130	−0.028
	(0.194)	(0.214)	(0.112)	(0.118)
Exclusion of voters	0.193*	0.234	−0.042	−0.076
	(0.113)	(0.161)	(0.134)	(0.185)
Illegal campaigning	0.036	0.019	0.102	0.096
	(0.084)	(0.158)	(0.107)	(0.101)
Observers excluded	0.567**	0.453	0.176	0.070
	(0.265)	(0.333)	(0.169)	(0.200)
Faulty ballot box	0.056	−0.129	−0.039	−0.108
	(0.142)	(0.297)	(0.124)	(0.124)
Secrecy violated	0.169**	0.104	0.086	0.079
	(0.066)	(0.092)	(0.054)	(0.057)
Illegal voting	0.216	0.426	0.050	0.100
	(0.228)	(0.347)	(0.165)	(0.191)
Other violations	0.872**	0.497	0.283	0.249
	(0.398)	(0.564)	(0.319)	(0.373)
Election FE	Y	Y	Y	Y

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: % Population with internet access; UR members per cap.

relate it to various types of election rigging for each subdivision. Table 1 presents the 32 coefficients from regressing the amount of each 8 types of election irregularity on 4 variants of the TVSC measure presented in the preceding section. In addition to that, the regressions also control for election specific trends and the incumbent party's strength proxied by party members per capita. The latter ensures validity of the TVSC which could also be driven by a party's advantages in mobilising voters and in turn increase the likelihood of reporting irregularities out of revenge. Finally, since fraud reports are likely to be incomplete and affected by technical impediments such as the lack of internet access, I also include the percentage of households with internet access into the regressions.

The results indicate that only the regional vote share of districts with a TVSC exceeding 1 is reliable. Most notably, it is strongly correlated with reports on those irregularities associated with fraud such as *improper counting* and *exclusion of observers*. Also other violations like *exclusion of voters*, *violation of secrecy*, and *other violations* seem to be correlated with the first indicator. The measure in column 2 produces similar but far less precise estimates. Indicators 3 and 4 which assume

Table 2: High TVSC and introduction of electronic ballot boxes, 2000-2012

	Vote share with TVSC ≥ 1			
	(1)	(2)	(3)	(4)
% Electronic ballot boxes	0.352	-0.685^{**}	-0.365^{***}	-0.366^{***}
	(0.263)	(0.296)	(0.117)	(0.115)
County FE	N	N	Y	Y
Election FE	N	Y	Y	Y
Controls	N	N	N	Y
Counties	71	71	71	71
Observations	497	497	497	497
R^2	0.003	0.105	0.762	0.762

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: log(GDP per capita)

a uniform distribution of fraud are only weakly and sometimes even negatively affecting the detection of irregularities. This makes the vote share of districts with a TVSC exceeding 1 the prime candidate for detecting election fraud in the Russian context. The inclusion of United Russia members per capita furthermore ensures reliability and that the results are not driven by other potential correlates of the TVSC.

Another desirable property for a good detector of rigging intensity is that it comoves with the opportunities for falsifying results. A significant decrease in such opportunities was marked by the start of electronic vote counting via optical scanners (KOIB) during the mid-2000s across Russia in about 5% of all voting stations. The introduction was staggered starting with the 2007 Duma elections which makes it unlikely to be correlated with other incentives for fraud. From official government documents I collected information on the numbers of KOIBs in each region across federal elections and calculated the percentage of voting stations equipped with such a device. Table 2 shows the results from regressing the vote share with TVSC$\geq$ 1 on the share of regions' voting stations equipped with a KOIB over the time period 2000 to 2012. Once election and region fixed effects are controlled for, one can see that having 1% more stations with electronic ballot boxes in a region reduces the share of votes from suspicious districts by 0.3%. The results of this test suggest again that the indicator is reliable and valid.

Admittedly, the share of suspicious votes based on TVSC$\geq$ 1 is not a perfect measure of rigging intensity but the findings above suggest that it is strongly correlated with what it is supposed to measure and that many potential concerns can

be ruled out. Unfortunately, there is no way to look at fraud reports before 2011 so some warranted doubt may still remain. In the following section I briefly present other indicators common in the analysis of rigged elections and check their explanatory strength in the case of Russia 2000 to 2012.

3.3 Alternative indicators

The preceding chapters have solely investigated quantitative tools for detecting turnout inflating types of rigging and the vast amount of evidence indeed points in the direction of this being the most widely used technique of manipulating the results in favour of a specific candidate (Filippov and Ordeshook, 1997; White, 2011; Enikolopov et al., 2013). Yet, anecdotal evidence from the Russian republics – Tatarstan, Ingushetia and Dagestan in particular – suggests that election results in some areas may not only be manipulated but entirely fabricated (Myagkov et al., 2009; Lukinova et al., 2011). In detecting this kind of fraud I follow the methodology of Beber and Scacco (2012) who rely on human preferences for specific numbers and biases in number generation. The main argument is that, under fairly generous assumptions, the final digit as well as the distance between the last and second-last digit of the vote count should follow a uniform distribution. In order to create alternative indicators of the share of suspicious votes, I adapt the methodology of Beber and Scacco to identify fraud at the district level and then aggregate this to the regional level using the share of affected votes as for the TVSC. In detail, I proceeded as follows: first, I calculated for each district the p-values of a Pearson's chi-squared test of uniform distribution of the last digit and the distance between last and second-last digits and repeated this procedure for both the reported valid votes as well as incumbent votes. In a second step, I obtained the share of votes in each region from districts where the hypothesis of uniform distribution could be rejected at a significance level of either 5 or 1%. This procedure has the advantage of being completely agnostic about the kind of bias, i.e. whether there is a bias towards fives in one district vs. eights in another, and only assumes whether votes for the incumbent or the number of valid votes/turnout were affected. The measures are furthermore also formed as ratios and therefore easily comparable to the indicators of suspicious votes presented above.

Table 3 repeats the analyses of table 1 for each of the 8 measures of fabricated votes. Interestingly, none of them has a significant positive effect on any reported election irregularity. If anything, some kinds of reports per capita are negatively affected by the share of votes from districts with numeric anomalies. Unusual digit distributions in the incumbent votes, for example, are associated with less reports on

Table 3: Reported irregularities 2011-2012 and numeric anomalies

Reports per 100k electorate	Effect of vote share with numeric anomalies for							
	Valid votes				Incumbent votes			
	last digit, with $p \leq$		Δ last 2 digits, with $p \leq$		last digit, with $p \leq$		Δ last 2 digits, with $p \leq$	
	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Improper counting	−0.303	0.379	0.305	0.388	−0.373	−0.489	0.625	−0.071
	(0.377)	(1.874)	(0.642)	(0.899)	(0.373)	(1.035)	(0.937)	(0.902)
Exclusion of voters	−0.295	3.647	2.136	−0.619	−0.274	−0.296	1.217	0.526
	(0.396)	(2.999)	(1.903)	(0.622)	(0.240)	(0.873)	(0.880)	(0.921)
Illegal campaigning	0.608	0.747	1.272	0.484	−0.473*	−0.147	−0.308	−0.045
	(0.525)	(1.276)	(0.947)	(1.140)	(0.258)	(0.839)	(0.403)	(0.520)
Observers excluded	0.002	3.157	−0.438	−0.788	−0.709	−0.351	1.570	−1.486
	(0.727)	(3.529)	(0.841)	(1.045)	(0.766)	(2.321)	(1.343)	(1.267)
Faulty ballot box	1.003	4.475	0.568	−0.890*	−0.969**	−0.918	−1.157*	−0.721
	(0.748)	(4.023)	(0.905)	(0.539)	(0.429)	(0.778)	(0.643)	(0.764)
Secrecy violated	−0.152	1.250	0.042	−0.505*	−0.072	0.053	0.360	0.257
	(0.178)	(1.252)	(0.215)	(0.300)	(0.145)	(0.461)	(0.340)	(0.423)
Illegal voting	0.653	12.040	0.300	−1.161	−1.254*	−0.757	0.404	0.055
	(1.078)	(8.138)	(0.922)	(0.711)	(0.675)	(1.249)	(1.165)	(2.180)
Other violations	2.057	15.814	1.201	−1.218	−0.543	−0.492	0.439	0.610
	(1.935)	(12.250)	(1.748)	(1.446)	(1.050)	(1.837)	(1.655)	(1.943)
Election FE	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: % Population with internet access

illegal campaigning, *faulty ballot boxes*, and *illegal voting*. Anomalies in the distance between the last two digits in the number of valid votes, on the other hand, significantly reduces also reports on *faulty ballot boxes* as well as the *violation of secrecy*. In particular, it does not seem to be correlated with *exclusion of observers*. One possible interpretation of these results is that fabricating votes may be a complement to turnout inflating methods which does apparently not require actions observable to the public and hence results in lower amounts of reported fraud. Voters' inability to detect and report entirely invented results makes it hard to verify the reliability of any fraud indicator built on numeric anomalies. The results in table 3 should thus also be regarded as a reminder that some fraud may still not be detected by the fraud indicator used in the remainder of this paper. Bearing this caveat in mind, I now turn towards the extent of potentially fraudulent election outcomes and its changes over time.

3.4 The evolution of suspicious results 2000–2012

Figure 3 plots for each election the share of votes with TVSC$\geq$ 1 in each Russian region. This helps understanding the variation in the main outcome variable and at the same time also shows how rigging changed at the extensive margin over the time period studied. While always present to some degree in few subdivisions, suspicious results started to take off during the March 2004 Presidential election from an average of 8 to 20%. During the 2007 Duma election it kept on rising to about 24-25% and remained roughly stable onwards. The distribution is skewed to the left resulting in median values below the corresponding means. Since the year 2003, however, also these gaps have been widening from 5 to about 10 percentage points indicating that intensity has surged disproportionately at the right tail.

Figure 4 zooms into the district level measures of rigging for the two elections before and after 2005. The plotted variable is a dummy whether the TVSC is exceeding unity in a given district and visualises how fraud evolved at the intensive margin across the Russian Federation. Already in 2003 suspicious districts are strongly concentrated in particular regions, most notably in the Republics of Tatarstan and Mordovia in Western Russia. In the March 2004 election suspicious votes start showing up in a number of formerly *clean* areas and additionally further regions start showing almost uniformly ballot counts with a TVSC$\geq$ 1 – Republics of Tuva in South-Central and Bashkortostan in the South-West. Other areas, particularly in the West and conflict-ridden South-Western Causcasus territory, are joining in during the next round of elections 2007/2008. This extreme concentration suggests

that time and region-specific characterstics could indeed be an important driver of prolific fraudulent election outcomes since 2000.

4 Data

4.1 Elections and fraud reports

The organisation of federal elections in Russia is roughly corresponding to its administrative divisions, both horizontally and vertically. The highest authority is the Central Election Commission (CIK), a permanent body whose members are nominated by the president, the second chamber *Council of Regions* and the *Duma* parliament. The CIK's main tasks are the coordination of the 83 Regional Election Commissions (IKS) and the organisation of the Federal elections. The IKS fulfils the same role as the CIK at the regional level and coordinates the territorial election commissions (TIKs). Unlike the CIK, its members are appointed by recommendation of the Federal government. The next administrative level below the region is the district (*rayon*). Like regions, the districts can vary considerably in size and population but unlike the former there may be several TIKs within the same rayon. This is especially often the case in larger cities or former *closed towns*. Over the period studied there were almost 3,000 TIKs. The members of these are permanent delegates by the regional executive, legislative and parties. The TIKs are therefore the level of electoral administration where the long arm of the central government starts to lose its power. At the lowest level, about 95,000 precinct election commissions (UIKs) are responsible for the local organisation of all elections and, most importantly, the vote counting. Unlike the other commissions, they are only formed one month before the elections and are nominated by the electorate. It remains unclear to what extent authorities can still exert control over the composition of the UIKs but their ad-hoc nature makes them unlikely to be the driving force behind organised large-scale fraud (OSCE, 2000, 2004a,b, 2012a,b).

All voting data used in this project comes from the organisation GOLOS, an independent Russian NGO concerned with election monitoring. The data covers each of presidential and parliamentary election since 2000 and reports results in absolute terms at the UIK level for the entire Russian Federation. The dataset also features official numbers on the electorate as well as valid and invalid votes required to calculate turnout in each precinct. I matched the results at the district/TIK level which is the main unit of observation in calculating the measures of electoral fraud explained in section 3. From GOLOS I also obtained direct indicators of election

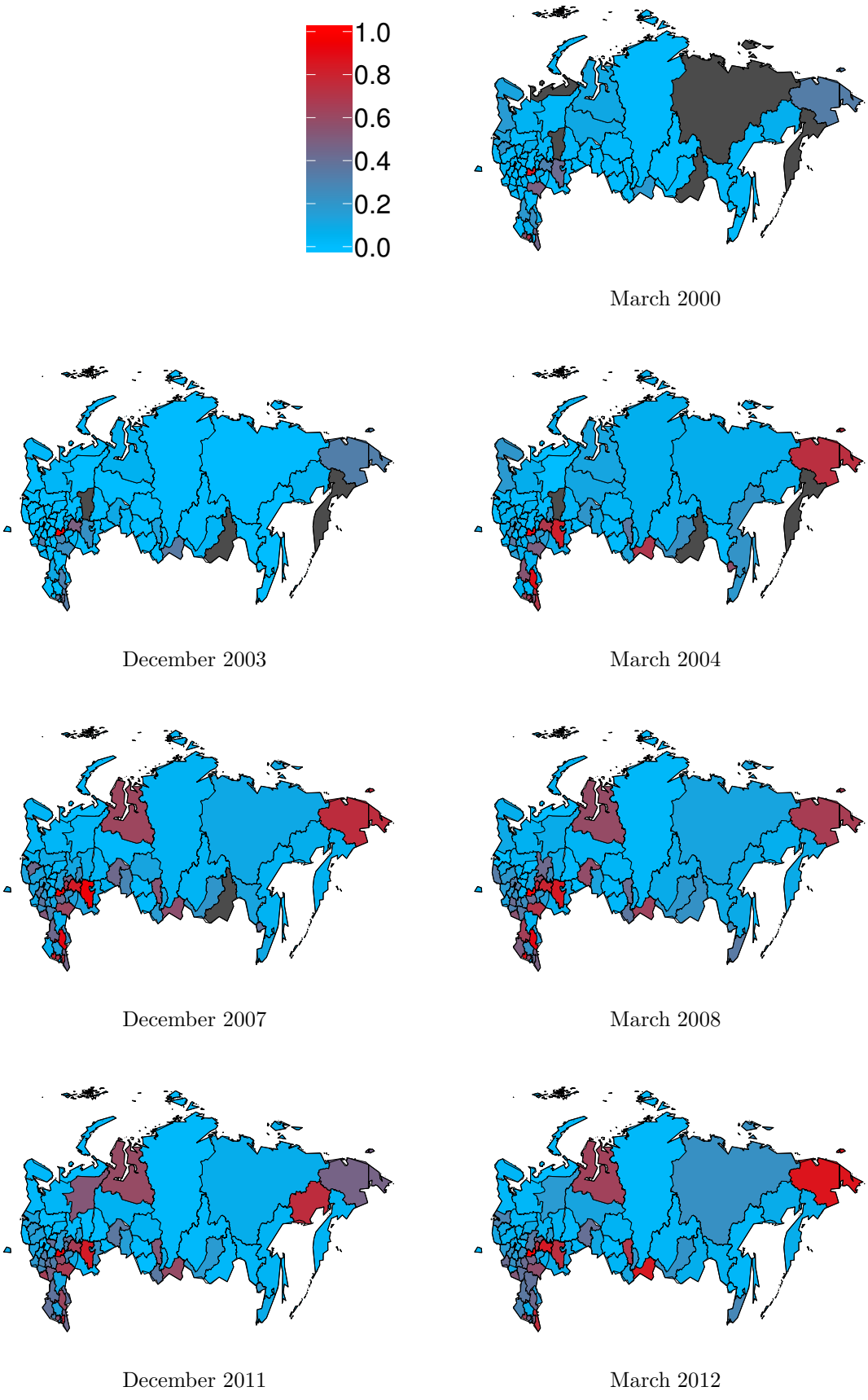
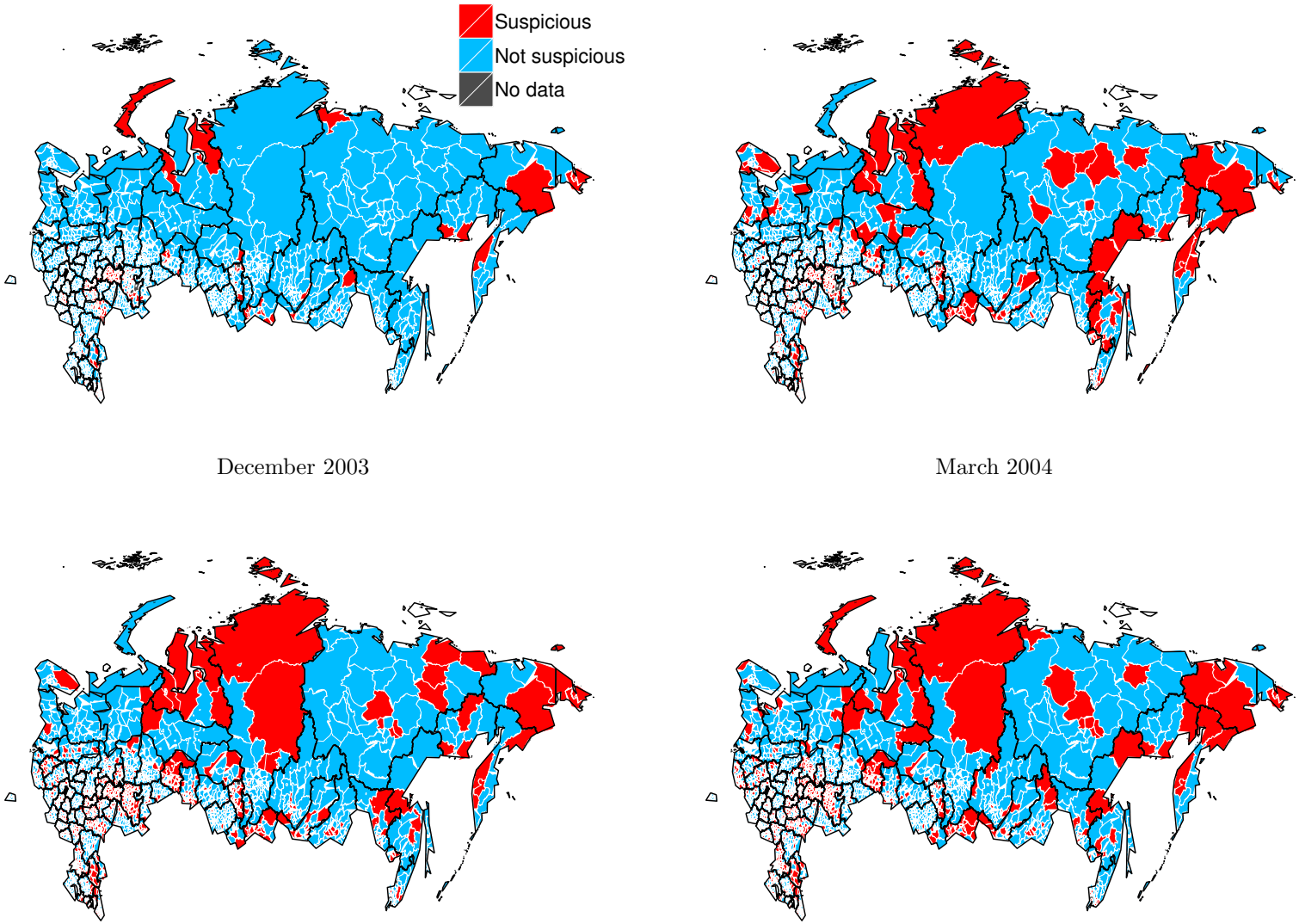


Figure 3: Share of suspicious votes before/after the abolition of governor elections in December 2004



December 2003 March 2004

December 2007 March 2008

FIGURE 4: SUSPICIOUS DISTRICTS BEFORE/AFTER THE ABOLITION OF GOVERNOR ELECTIONS IN DECEMBER 2004

rigging. During the 2011 and 2012 elections, the association ran the *karta narusheniy* (map of violations) project which provided a platform for citizens to anonymously report incidents of fraud in federal, regional, and local elections and send detailed reports of observed electoral law violations via phone, internet, and text message. In addition to that, the users could also give information whether the action happened during campaigning or on election day and which type of fraud had taken place. Of particular interest for this research project are the categories *distortion of results* and *exclusion of observers, committee members, or media*. Lastly, the observers could also provide information on the location where the action was witnessed. I used this information to match each *violations reports* to a specific district which was possible for about 80% of all 12,800 reports (Duma 2011: 6,200/7,800; President 2012: 4,000/5,000). The analysis also draws on a number of other variables which are explained in the following.

4.2 Socio-economic data, further variables, and sample

According to reports, the main goal of the central government apart from delivering high vote results and legitimacy is the maintaining of social and economic stability. What is far less clear, however, is how such stability is evaluated by Russia's leadership. Obvious indicators relate to the economic performance of a region, especially those also related to the well-being of the population. I therefore obtained panel data from the Russian Federal State Statistics Service (GKS) on regions' log GDP per capita and unemployment rate. Assuming that governors are inferring the electorate's support for the incumbent party and presidential candidate based on their economic performance over the term time, I calculate 4-year changes in both variables. In order to get a single indicator of economic prosperity, the changes are then aggregated to a single index using principal component analysis. In order to assess the index' validity as a proxy for political support, I use data from the *Public Opinion Foundation* compiled in Reuter and Robertson (2012). This yearly measure is the % of survey respondents in 68 Russian regions answering positively to the question whether their governor is doing a good job or not. From figure 5 one can see that the index is positively correlated with approval of a region's head. A simple OLS regression with region clustered errors and region fixed effects yields a positive and significant coefficient with a t-statistic of 5.4. Finally, in section 3 I make use of panel data on regions' population with internet access which was also obtained from GKS.

The switch from an *elected* governor to a *handpicked* governor is coded in the main treatment variable $HandpickedGovernor_{it}$. It has value 1 if the governor ruling

Figure 5: Economic performance and governor popularity

region i at time t started his term after the 12$^{\text{th}}$ of December 2004 – and therefore had to be selected by the Russian president – or 0 otherwise. The precise dates of when a region's head entered office were retrieved from the website *rulers.org*. Another important piece of information are numbers on the members of the ruling party *United Russia* by region. This information was gathered from reports of Russian Federal Ministry of Justice for the years 2009 until 2012 and turned into a per capita measure using yearly population counts at the region level from GKS.

The final panel dataset covers 71 out Russia's 83 regions over all 7 federal elections during the period 2000 to 2012. During this time there were 5 mergers between 2 or 3 regions which reduced the initial amount of 89 subdivisions to 83. Such mergers are likely to fundamentally change the power structure of a governor and make it difficult to compare the new units especially since all mergers took place after the abolishment of governor elections. For this reason I excluded all 10 regions affected by a merger. In addition to that, Chechnya lacked information on economic outcomes until the early 2000s and the Republic of Sakha didn't provide precinct-level voting data in 2000. Both regions were hence dropped from the dataset. Summary statistics of the final sample are reported in table 4.

Table 4: Descriptive statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
Share of ballot w/ TVSC≥1	497	0.21	0.24	0.00	1.00
% Incumbent vote	497	0.59	0.16	0.25	0.99
% Turnout	497	0.65	0.11	0.44	0.98
Electorate in 100,000	497	13.60	12.35	0.34	73.10
Reports per 100k electorate on...					
Improper counting	142	0.36	0.49	0.00	2.56
Exclusion of voters	142	0.26	0.58	0.00	6.31
Illegal cmapaigning	142	0.12	0.32	0.00	2.92
Observers excluded	142	0.55	0.79	0.00	5.07
Faulty ballot box	142	0.24	0.42	0.00	2.92
Secrecy violated	142	0.12	0.16	0.00	0.93
Illegal voting	142	0.45	0.72	0.00	5.59
Other violations	142	0.87	1.11	0.00	8.86
% Pop. w/ internet	142	0.41	0.14	0.01	0.72
United Russia members p.c.	142	16.47	8.95	5.61	53.45
Handpicked governor	497	0.30	0.46	0	1
Handpicked governor b/w 2004/2007	497	0.34	0.47	0	1
Population in 100,000	497	17.99	17.17	0.51	118.43
4-year Δ Economy	497	0.02	1.14	−2.56	4.06
log(GDP p.c.)	497	11.37	0.91	8.75	14.06
4-year Δ log(GDP p.c.)	497	0.86	0.29	0.17	1.67
Unemployment rate	497	0.09	0.06	0.01	0.57
4-year Δ Unemployment rate	497	−0.02	0.04	−0.18	0.12

Notes: The unit of observation is one of the 71 regions in the sample at election t. Non-voting data available over several periods is linearly interpolated to the time of the election. Variables provided at the cross-sectional level only (i.e. with only 71 observations) are reported accordingly and used in the analysis by interacting them with either a post-2004 dummy or election fixed effects.

5 Empirical analysis

5.1 Identification

The main predictions of section 2.2 are that handpicked governors 1) have in general less incentives to rig elections since loyalty is assured and 2) they use rigging to compensate expected lower election results. One would therefore expect that in regions who had a handpicked governor, overall fraud levels decrease but that social instability could drive them up again. I model this mechanism in a difference-in-differences specification analysing the changing effect of economic performance on

the extent of election fraud. The treatment in this setup is the forced change to a governor chosen by the central administration:

$$\begin{aligned} ShareSuspicious_{it} = \alpha + \gamma_i + \lambda_t + \boldsymbol{\mu}\, \boldsymbol{X}_{it} \\ + \beta\, HandpickedGovernor_{it} + \mu\, \Delta Economy_{it} \\ + \theta\, HandpickedGovernor_{it} \times \Delta Economy_{it} + \epsilon_{it} \end{aligned} \tag{1}$$

The addition of election-specific and region fixed effects λ_t and γ_i restricts the focus only to variation in suspicious votes within regions off any election specific trend. Election FEs account for the strong upward shift in suspicious votes over time that is by construction correlated with the arrival of handpicked governors after 2005. Area-specific fixed effects further control for permanently strong political machines originating from times of the Soviet Union (Hale, 2003). Further controls $\boldsymbol{X}_{it}$ are including regions' aggregate democracy rating between 1991 and 2001 interacted with election fixed effects. The standard errors are clustered at the region level to account for autocorrelation of region-specific unobservables which may downward bias conventional robust estimates of the residuals' variance. In order to consistently estimate the effect of having a handpicked governor and its interaction with economic performance on the share of suspicious votes, one needs to assure that treated regions did not systematically differ from non-treated ones and that replacement is not correlated with other simultaneous changes in the respective areas. The first assumption is casually checked in figure 6 which displays the mean share of potentially fraudulent votes over time for regions with and without a replacement of governor between 2004 and 2008. As can be seen, the two groups follow roughly similar trends before the new law, even during the first major increase of suspicious votes in the presidential elections of March 2004. In the following elections, the patterns start diverging with *not replaced* regions displaying notably higher levels of suspicious votes. After an initial peak in 2008 with a difference in averages of 10%, *replaced* regions remain about 5% below the level of their counterparts.

Unlike the common trends assumption, the absence of confounding events is not straightforward to check. Replacing a region's leader may not only mean a change in loyalty but also in many other factors potentially correlated with rigging incentives. Two such confounders may be that the central government was targeting either particularly unpopular governors or particularly unsuccessful ones. Despite lacking pre-treatment data, figure 7 gives the hint that, if anything, handpicked governors were less popular than elected ones at the beginning but caught up over

Figure 6: Average suspicious vote share before/after introduction of governor appointments

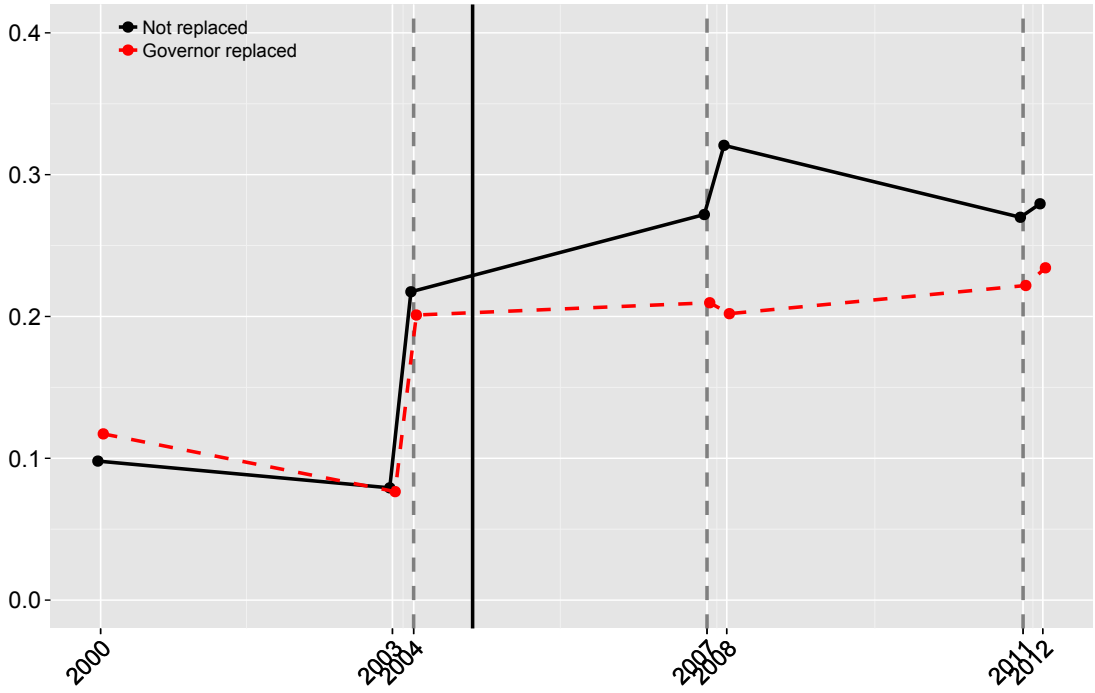


Figure 7: Average popularity of governors

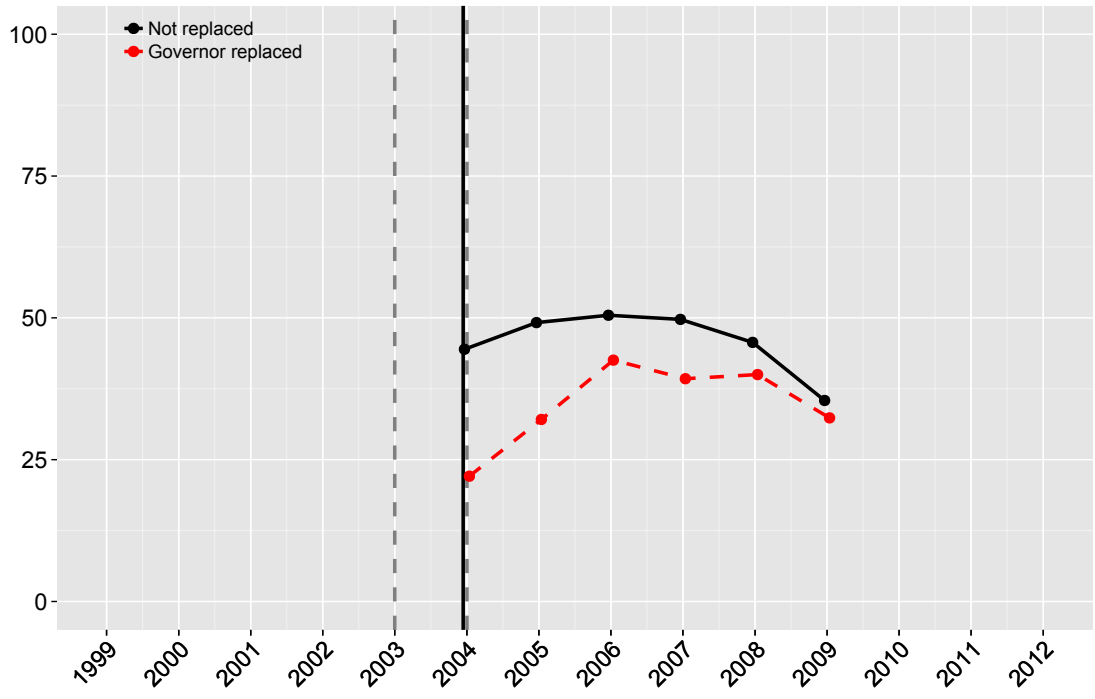


Figure 8: Average incumbent vote share before/after introduction of governor appointments

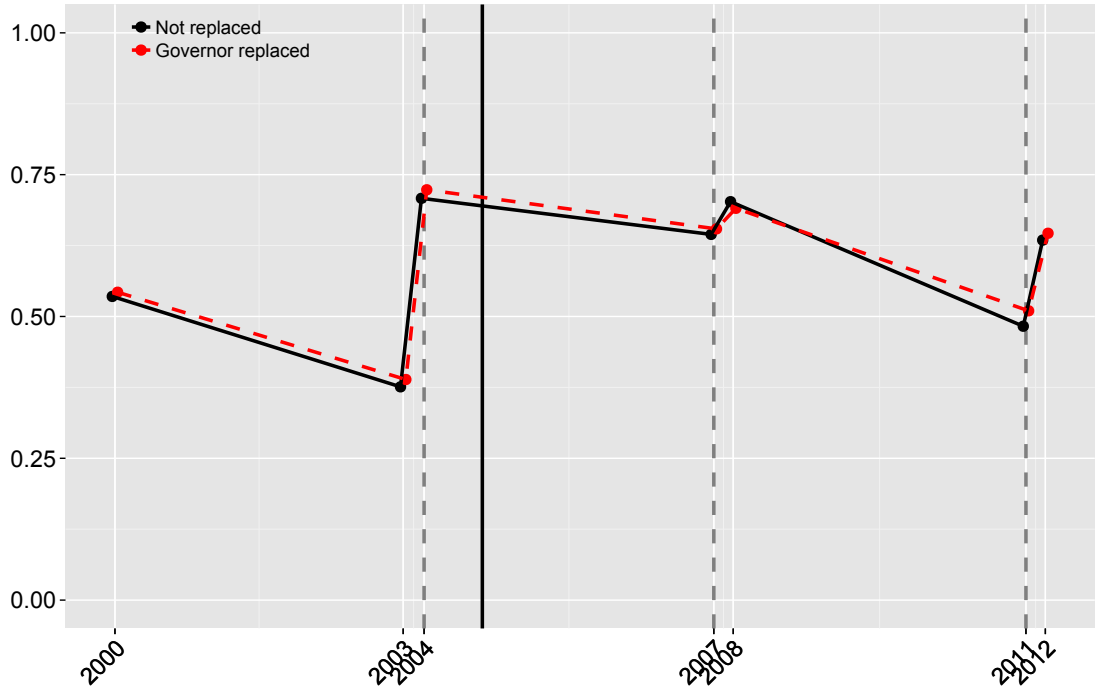
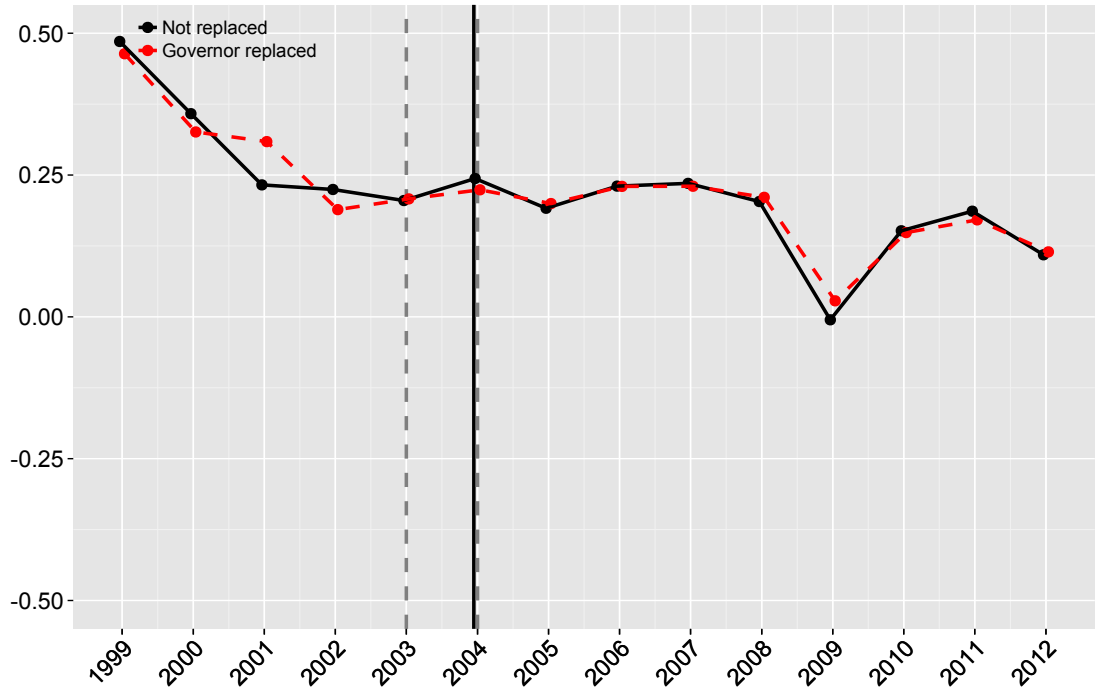


Figure 9: Average yearly GDP p.c. growth before/after introduction of governor appointments



time. Given this, one would expect that handpicked governors were to engage in *more* fraud than old ones since their popularity does not allow them to turn voters in favour of the central government. Figure 8 also highlights that incumbent votes in federal elections were at a similar level and that replacement does not appear to have been a punishment for pre-2005 election results. The fact that incumbent votes do not respond to replacement after 2005 is striking at first but is not informative about governors' popularity given the changing intensity of fraud used to produce those results. A more reliable check whether success was a selection criterion or a confounder is provided in figure 9 which compares the average yearly growth in regional GDP per capita across the two subgroups. Again, there is no evidence for substantial differences across regions before and after the appointment of governors. It does not appear that a new region's head spurred economic growth which could have led to lower need to rig federal elections.

5.2 Baseline results

Table 5 reports the main results of the difference-in-differences estimation. As can be seen from the first 2 columns, *HandpickedGovernor* is positively correlated with potentially fraudulent elections even when controlling for time-invariant regional characteristics. This, however, is due to the general rise in replaced governors and fraud over time which leads to a notable upward bias. Once election fixed effects are accounted for in column 3 the coefficient flips sign but remains highly significant. According to the estimates, having a handpicked governor reduces suspicious votes by 8.5%. This is even larger than the 5% difference observed in the raw data in figure 6 and equivalent to a third of a standard deviation or moving from the median to the 25^{th} percentile of the distribution. The result tentatively confirms the theoretical prediction that if the government selects its own candidate, the need to signal his loyalty by rigging elections is substantially diminished.

The coefficient remains virtually unaffected by the inclusion of control variables and 4-year economic performance in the next two specifications. Especially column 5 is reassuring that replacement is not endogenous to a leader's performance and that socio-economic development prior to the election has on average no effect on suspicious votes. The final specification allows this effect to differ for regions with a handpicked governor and shows that in this case, economic performance significantly lowers the share of suspicious votes in federal elections. In other words, the negative correlation between having an appointed governor and potentially fraudulent votes turns even more negative when the economy has done well over the last four years. Given that the minimum value of *4-year* Δ *Economy* is -2.56, handpicked governors

Table 5: Difference-in-Differences Results

	Share of suspicious votes					
	(1)	(2)	(3)	(4)	(5)	(6)
Handpicked Governor	0.037	0.064***	−0.086***	−0.086***	−0.086***	−0.108***
	(0.027)	(0.017)	(0.027)	(0.027)	(0.027)	(0.029)
4-year Δ Economy					0.005	0.016
					(0.013)	(0.014)
Handpicked Governor × 4-year Δ Economy						−0.047**
						(0.021)
Region FE	N	Y	Y	Y	Y	Y
Election FE	N	N	Y	Y	Y	Y
Controls	N	N	N	Y	Y	Y
Regions	71	71	71	71	71	71
Observations	497	497	497	497	497	497
R^2	0.005	0.674	0.770	0.787	0.787	0.790

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: $\text{Democracy}_{1991-2001}$ × Election FE

deliver similarly suspicious results *only* when the region's economic development was extremely bad. This finding is consistent with hypothesis 2 in section 2.2 that even selected officials may use rigging to signal competence in the absence of political success.

5.3 Sensitivity and robustness checks

As a first test for the stability of the baseline estimates, I re-estimate the original model including additional fixed effect specifications. The two setups I am using are time shifters for each of the 8 Federal Districts and region-specific linear time trends. Despite looking similar at first, their inclusion serves different purposes. Federal Districts were created in 2000 by President Putin as an intermediary subdivision between the federal government and the regions and cover between 6 to 18 of these. The corresponding *plenipotentiaries* are directly appointed by the president and were used to tighten control over territories' leaders (Hill, 2012). A particularly ambitious district leader could therefore replace corrupt governors and simultaneously disincentivise ballot rigging or introduce manipulation techniques that the TVSC cannot capture which would give similar results to the ones in section 5.2. Region-specific linear time trends, on the other hand, provide a test whether the effect could be driven by diverging trends in fraud between treated and non-treated regardless of governor replacements. Table 6 depicts the baseline esti-

Table 6: Baseline results and different FE specifications

	Share of suspicious votes			
	(1)	(2)	(3)	(4)
Handpicked Governor	−0.108***	−0.096***	−0.079**	−0.082**
	(0.029)	(0.030)	(0.035)	(0.037)
4-year Δ Economy	0.016	0.008	0.011	0.004
	(0.014)	(0.015)	(0.016)	(0.018)
Handpicked Governor × 4-year Δ Economy	−0.047**	−0.035	−0.049**	−0.050*
	(0.021)	(0.023)	(0.025)	(0.027)
Region FE	Y	Y	Y	Y
Election FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Election × Fed.Distr. FE	N	Y	N	Y
Region FE × t	N	N	Y	Y
Regions	71	71	71	71
Observations	497	497	497	497
R^2	0.790	0.821	0.853	0.873

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: $\text{Democracy}_{1991-2001}$ × Election FE

mates and their sensitivity to including the two additional fixed effect specifications jointly and by themselves. Reassuringly, the estimates for *HandpickedGovernor* and *HandpickedGovernor×4-year Δ Economy* are not changing substantially in magnitude and remain significant. The level treatment effect increases to about -0.08 in column 3 and 4, suggesting that the baseline estimate was slightly biased downwards due to diverging trends. The interaction with economic performance is less precisely estimated but remains almost identical even in the most flexible fixed effects setup in the final column. Overall, it seems unlikely that characteristics at the federal district level or diverging patterns in election fraud are driving the baseline results.

Another important question is the sensitivity of the baseline findings for *HandpickedGovernor × 4-year Δ Economy* to alternative definitions of economic performance. In order to evaluate this, I investigate in table 7 the sensitivity of the baseline results to varying time-horizons of economic performance as well as the individual components of the index. Column 2 to 4 show that using the PCA of changes in unemployment and log GDP per capita over time horizons closer to the election data does not change the treatment effect. The coefficient on the interaction effect *HandpickedGovernor×4-year Δ Economy*, however, halves in magnitude and loses its significance. The choice of a 4-year horizon thus seems to be a crucial choice for finding evidence on the competence mechanism. Finally, specifications 5

Table 7: Sensitivity of results to definition of economic performance

Outcome variable	PCA of Δ Unemployment rate and log(GDP p.c.) over				4-year Δ of	
	4 years (baseline)	3 years	2 years	1 year	Unemployment rate	log(GDP p.c.)
	(1)	(2)	(3)	(4)	(5)	(6)
Handpicked Governor	−0.108***	−0.093***	−0.098***	−0.093***	−0.071***	−0.001
	(0.029)	(0.028)	(0.026)	(0.027)	(0.026)	(0.065)
Δ Economy	0.016	0.010	0.017	0.023	−0.380	0.050
	(0.014)	(0.013)	(0.012)	(0.015)	(0.356)	(0.061)
Handpicked Governor × Δ Economy	−0.047**	−0.019	−0.028	−0.023	1.287**	−0.120
	(0.021)	(0.024)	(0.019)	(0.020)	(0.587)	(0.085)
Region FE	Y	Y	Y	Y	Y	Y
Election FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Regions	71	71	71	71	71	71
Observations	497	497	497	497	497	497
R^2	0.790	0.787	0.789	0.790	0.790	0.788

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: $\text{Democracy}_{1991-2001}$ × Election FE

and 6 reveal that 4-year changes in unemployment are probably the main driver of the baseline results, while changes in log GDP per capita over the same time period per se do not allow reproducing these findings.

5.4 Placebo tests and related outcomes

The plausibility of the common trends assumption can also be tested through a placebo experiment in which the treatment period is moved forward by several time periods. In doing so, one can test whether there was a significant level difference between treatment and control group prior to the selection of a new governor which may have not been captured by the region-specific time trends in the preceding section. For instance, a new law in 2001 made it possible to prosecute governors for criminal activities (Sharafutdinova, 2010). If under-performing, to-be-removed governors were decreasing election fraud in response to this law or any other policy change or if fraud reduction was in fact an anticipatory behaviour, this may still yield results similar to the baseline. Table 8 reports the initial estimates along with two further specifications looking only at pre-2005 data where the replacement of a governor is moved either one or two elections forward. As can be seen from column 2 and 3, this manipulation of the treatment variable halves the corresponding point estimate of *HandpickedGovernor* and leaves it insignificant. The interaction with *4-year* Δ *Economy* loses its significance in specification 3 but flips sign and remains highly significant in the second case. This finding can be explained by the fact that the 2004 surge in rigging was particularly strong in the Northern Caucasus regions such as Dagestan and North Ossetia whose governors were among the first ones to be replaced and at the same time saw their economies recovering after the nearby Second Chechen War 1999 to 2000. Once the 6 regions of the Northern Caucasian Federal district are omitted, the coefficient drops and becomes insignificant (see table A1).

One implicit assumption of this study is that election fraud in Russia between 2000 and 2012 has been mainly turnout increasing. This means that additional votes for the advantaged candidate are generated through unused ballot sheets or biased mobilisation of voters rather than stealing votes from other candidates or parties. Even the vote share of the incumbent may not be affected, given reports on specific *vote targets* which could be achieved by legal and non-legal means (White, 2011). Hence, one would expect turnout but not other parties to respond to having a handpicked governor which offers another insightful placebo test. I therefore re-estimate equation 1 using turnout and vote shares of the incumbent and other parties (Communist, Ultranational, and Democratic) as outcomes. The findings

Table 8: Results from placebo tests on pre-treatment data

Treatment	Handpicked Governor$_t$	Handpicked Governor$_{t+1}$	Handpicked Governor$_{t+2}$
	(1)	(2)	(3)
Handpicked Governor	−0.090***	−0.062	−0.059
	(0.028)	(0.038)	(0.049)
4-year Δ Economy	0.009	−0.009	−0.003
	(0.013)	(0.017)	(0.018)
Handpicked Governor × 4-year Δ Economy	−0.031	0.067**	0.020
	(0.021)	(0.030)	(0.021)
Region FE	Y	Y	Y
Election FE	Y	Y	Y
Controls	Y	Y	Y
Sample	2000–2012	2000–2004	2000–2004
Regions	69	69	69
Observations	483	207	207
R^2	0.790	0.849	0.847

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: Democracy$_{1991-2001}$ × Election FE

presented in table 8 are roughly in line with the assumption of turnout increasing election rigging. Turnout in column 2 is decreasing by 2% on average in regions with an appointed governor, significant at the 10% level. The marginal effect turns even more negative depending on economic performance but is not statistically significant. The effect on incumbent vote share in specification 3 yields qualitatively similar but insignificant results. For the remaining parties, both the *HandpickedGovernor* and *HandpickedGovernor*×*4-year* Δ *Economy* coefficients are far smaller and never significant. The coefficients on *4-year* Δ *Economy* indicate that, even in the presence of large-scale fraud, voters may hold the central government accountable to some extent and give their votes to opposition parties if their economic situation worsens. An effect of having a handpicked governor on vote shares of the incumbent or any other party can, however, not be found.

In sum, the results of section 5.2 have proven stable throughout a number of robustness and falsification checks. The interaction with regional economic growth has, however, turned out slightly less stable and appears to depend strongly on the time dimension used to construct the principal components.

Table 9: Results for different election outcomes

Dependent variable	Share suspicious votes	% Turnout	% Incumbent	% Communist	% Ultra-national (LDPR)	% Democratic (Yabloko)
	(1)	(2)	(3)	(4)	(5)	(6)
Handpicked Governor	−0.108***	−0.020*	−0.013	0.006	−0.005	0.004
	(0.029)	(0.011)	(0.013)	(0.007)	(0.003)	(0.002)
4-year Δ Economy	0.016	0.005	0.022***	−0.013***	−0.002*	0.000
	(0.014)	(0.004)	(0.005)	(0.003)	(0.001)	(0.002)
Handpicked Governor × 4-year Δ Economy	−0.047**	−0.012	−0.010	0.004	−0.003	0.000
	(0.021)	(0.013)	(0.012)	(0.006)	(0.003)	(0.003)
Region FE	Y	Y	Y	Y	Y	Y
Election FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Regions	71	71	71	71	71	71
Observations	497	497	497	497	497	497
R^2	0.790	0.825	0.894	0.857	0.898	0.860

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: $\text{Democracy}_{1991-2001}$ × Election FE

6 Conclusion

In this paper, I study the importance of incentive structures of local officials on the dynamics of election fraud in Russia. I exploit a radical law change in December 2004 which allowed the central government to remove governors without any constraints and thus created strong motivation for the latter to use rigging in order to stay in office. Hypotheses from a simple conceptual framework predict that governors handpicked by the central government have less need to engage in fraud than elected (and not yet replaced) ones since their loyalty is assured. Handpicked governors on the other hand are assumed to respond with rigging in the face a bad economic performance and lower expected votes. The paper develops and extensively tests a new indicator of electoral fraud for Russian regions between 2000 and 2012 which is created from a unique micro-level dataset of election results at the voting station level.

The effect of having a handpicked governor on the share of suspicious votes in a region is estimated using a differences-in-differences estimation. The baseline results support the hypotheses and showed that regions with a handpicked governor have on average 10% less suspicious votes than those with elected ones. Furthermore, also the interaction of the treatment variable with economic performance is negative and indicates that in the case of a very bad economic performance the extent of fraud by handpicked governors would be equal to that of elected ones. In this sense, loyalty and competence can be regarded as complementary. Both effects were highly significant and passes several robustness checks concerning the validity of the common trends assumption and placebo treatments. While the share of suspicious votes is affected by the law change and handpicked governors, I also show that election outcomes of the incumbent as well as other parties did not respond. In other words, the incentive structures of governors does not change the results of elections, but only the way they are generated.

Despite focussing only on the case of Russia, the findings provide interesting insights into the functioning of competitive authoritarian systems in general. Unlike in a totalitarian system, elections can actually still function as an arena for political competition, albeit only among lower-tier officials. Contrary to common knowledge, I show conceptually and empirically that rigging is far less common among the dictator's cronies due to their certain loyalty. From a policy perspective, this means that advocating for the co-optation of non-cadres to government positions in an authoritarian regime may actually have counter-productive results and lead to higher

levels of election fraud. Also the fact that bad economic performance induces even higher levels of election fraud should be born in mind when deciding about sanctions.

References

Beber, B. and A. Scacco (2012). What the numbers say: A digit-based test for election fraud. *Political Analysis 20*(2), 211–234.

Cantu, F. and S. M. Saiegh (2011). Fraudulent democracy? an analysis of argentina's infamous decade using supervised machine learning. *Political Analysis 19*(4), 409–433.

Chaisty, P. (2012). The federal assembly and the power vertical. In G. J. Gill and J. Young (Eds.), *Routledge handbook of Russian politics & society*, pp. 92–101. London: Routledge.

Chaykovskaya, E. (12/12/2011). Vologda region governor quits after poor election result. *The Moscow News*.

Egorov, G. and K. Sonin (2011). Dictators and their viziers: Endogenizing the loyalty–competence trade-off. *Journal of the European Economic Association 9*(5), 903–930.

Enikolopov, R., V. Korovkin, M. Petrova, K. Sonin, and A. Zakharov (2013). Field experiment estimate of electoral fraud in russian parliamentary elections. *Proceedings of the National Academy of Sciences 110*(2), 448–452.

Filippov, M. and P. C. Ordeshook (1997). Who stole what in russia's december 1993 elections. *Demokratizasiya 5*(1), 36–52.

Gandhi, J. and E. Lust-Okar (2009). Elections under authoritarianism. *Annual Review of Political Science 12*(1), 403–422.

Goode, J. P. (2007). The puzzle of putin's gubernatorial appointments. *Europe-Asia Studies 59*(3), 365–399.

Hale, H. (2003). Explaining machine politics in russia's regions: Economy, ethnicity, and legacy. *Post-Soviet Affairs 19*(3), 228–263.

Haraszti, M. (Ed.) (2004). *Report on Russian media coverage of the Beslan tragedy: Access to information and journalists' working conditions*. The Representative on Freedom of the Media.

Hill, R. J. (2012). The putin era. In G. J. Gill and J. Young (Eds.), *Routledge handbook of Russian politics & society*, pp. 13–26. London: Routledge.

Ichino, N. and M. Schündeln (2012). Deterring or displacing electoral irregularities? spillover effects of observers in a randomized field experiment in ghana. *Journal of Politics 74*(01), 292–307.

Kalinin, K. and W. R. Mebane (2011). Understanding electoral frauds through evolution of russian federalism: from "bargaining loyalty" to "signaling loyalty". mimeo.

Lukinova, E., M. Myagkov, and P. C. Ordeshook (2011). Metastasised fraud in russia's 2008 presidential election. *Europe-Asia Studies 63*(4), 603–621.

Martinez-Bravo, M. (2014). The role of local officials in new democracies: Evidence from indonesia. *American Economic Review 104*(4), 1244–1287.

Moraski, B. J. and W. M. Reisinger (2009). Regional voting in russia's federal elections and changing regional deference to the kremlin. Prepared for delivery at the 67th annual national conference of the midwest political science association, chicago, il, april 2-5, 2009.

Moraski, B. J. and W. M. Reisinger (2013). Deference or governance? a survival analysis of russia's governors under presidential control. In W. M. Reisinger (Ed.), *Russia's Regions and Comparative Subnational Politics*, pp. 40–62. Routledge.

Myagkov, M. G., P. C. Ordeshook, and D. Shakin (2009). *The forensics of election fraud: Russia and Ukraine*. Cambridge: Cambridge University Press.

OSCE (2000). *Russian Federation, Presidential Election, 26 March 2000*. OSCE/ODIHR Election Observation Mission Final Report. Warsaw.

OSCE (2004a). *Russian Federation, Elections to the State Duma, 7 December 2003*. OSCE/ODIHR Election Observation Mission Final Report. Warsaw.

OSCE (2004b). *Russian Federation, Presidential Election, 14 March 2004*. OSCE/ODIHR Election Observation Mission Final Report. Warsaw.

OSCE (2012a). *Russian Federation, Elections to the State Duma, 4 December 2011*. OSCE/ODIHR Election Observation Mission Final Report. Warsaw.

OSCE (2012b). *Russian Federation, Presidential Election, 4 March 2012*. OSCE/ODIHR Election Observation Mission Final Report. Warsaw.

Reuter, O. J. and G. B. Robertson (2012). Subnational appointments in authoritarian regimes: Evidence from russian gubernatorial appointments. *Journal of Politics 74*(4), 1023–1037.

Rochlitz, M. (2013). Elections vs. appointments: Comparing incentive patterns for russian governors under putin. mimeo.

Sakwa, R. (2008). *Russian politics and society* (4 ed.). London: Routledge.

Sharafutdinova, G. (2010). Subnational governance in russia: How putin changed the contract with his agents and the problems it created for medvedev. *Publius: The Journal of Federalism 40*(4), 672–696.

Slider, D. (2012). Regional governance. In G. J. Gill and J. Young (Eds.), *Routledge handbook of Russian politics & society*, pp. 153–163. London: Routledge.

Szakonyi, D. (2012). You're fired!: Identifying electoral accountability in a competitive authoritarian regime. mimeo.

White, S. (2011). Elections russian-style. *Europe-Asia Studies 63*(4), 531–556.

Ziblatt, D. (2009). Shaping democratic practice and the causes of electoral fraud: The case of nineteenth-century germany. *American Political Science Review 103*(01), 1.

A Tables

Table A1: Results from placebo tests on pre-treatment data (omitting Northern Caucasus district)

Treatment	Handpicked Governor$_t$	Handpicked Governor$_{t+1}$	Handpicked Governor$_{t+2}$
	(1)	(2)	(3)
Handpicked Governor	-0.069^{**}	-0.050	-0.044
	(0.027)	(0.041)	(0.048)
4-year Δ Economy	0.021	-0.002	0.003
	(0.014)	(0.020)	(0.021)
Handpicked Governor $\times$ 4-year Δ Economy	-0.013	0.052	0.009
	(0.020)	(0.036)	(0.023)
Region FE	Y	Y	Y
Election FE	Y	Y	Y
Controls	Y	Y	Y
Sample	2000–2012	2000–2004	2000–2004
Regions	65	65	65
Observations	455	195	195
R^2	0.801	0.867	0.866

Notes: Standard errors clustered at the region level in parantheses, *p<0.1; **p<0.05; ***p<0.01; Controls: Democracy$_{1991-2001}$ $\times$ Election FE